Comparative Genomics of Vancomycin-Resistant Staphylococcus aureus Strains and Their Positions within the Clade Most Commonly Associated with Methicillin-Resistant S. aureus Hospital-Acquired Infection in the United States

Comparative Genomics of Vancomycin-Resistant *Staphylococcus aureus* Strains and Their Positions within the Clade Most Commonly Associated with Methicillin-Resistant *S. aureus* Hospital-Acquired Infection in the United States

Veronica N. Kos,[a,b] Christopher A. Desjardins,[b] Allison Griggs,[b] Gustavo Cerqueira,[b] Andries Van Tonder,[c] Matthew T. G. Holden,[c] Paul Godfrey,[b] Kelli L. Palmer,[a] Kip Bodi,[d] Emmanuel F. Mongodin,[e] Jennifer Wortman,[b] Michael Feldgarden,[b] Trevor Lawley,[c] Steven R. Gill,[f] Brian J. Haas,[b] Bruce Birren,[b] and Michael S. Gilmore[a,b]

Departments of Ophthalmology and of Microbiology and Immunobiology, Harvard Microbial Sciences Initiative, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts, USA[a]; The Broad Institute, Cambridge, Massachusetts, USA[b]; The Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom[c]; Tufts University School of Medicine, Boston, Massachusetts, USA[d]; Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, USA[e]; and School of Medicine and Dentistry, University of Rochester, Rochester, New York, USA[f]

ABSTRACT **Methicillin-resistant *Staphylococcus aureus* (MRSA) strains are leading causes of hospital-acquired infections in the United States, and clonal cluster 5 (CC5) is the predominant lineage responsible for these infections. Since 2002, there have been 12 cases of vancomycin-resistant *S. aureus* (VRSA) infection in the United States—all CC5 strains. To understand this genetic background and what distinguishes it from other lineages, we generated and analyzed high-quality draft genome sequences for all available VRSA strains. Sequence comparisons show unambiguously that each strain independently acquired Tn*1546* and that all VRSA strains last shared a common ancestor over 50 years ago, well before the occurrence of vancomycin resistance in this species. In contrast to existing hypotheses on what predisposes this lineage to acquire Tn*1546*, the barrier posed by restriction systems appears to be intact in most VRSA strains. However, VRSA (and other CC5) strains were found to possess a constellation of traits that appears to be optimized for proliferation in precisely the types of polymicrobic infection where transfer could occur. They lack a bacteriocin operon that would be predicted to limit the occurrence of non-CC5 strains in mixed infection and harbor a cluster of unique superantigens and lipoproteins to confound host immunity. A frameshift in *dprA*, which in other microbes influences uptake of foreign DNA, may also make this lineage conducive to foreign DNA acquisition.**

IMPORTANCE **Invasive methicillin-resistant *Staphylococcus aureus* (MRSA) infection now ranks among the leading causes of death in the United States. Vancomycin is a key last-line bactericidal drug for treating these infections. However, since 2002, vancomycin resistance has entered this species. Of the now 12 cases of vancomycin-resistant *S. aureus* (VRSA), each was believed to represent a new acquisition of the vancomycin-resistant transposon Tn*1546* from enterococcal donors. All acquisitions of Tn*1546* so far have occurred in MRSA strains of the clonal cluster 5 genetic background, the most common hospital lineage causing hospital-acquired MRSA infection. To understand the nature of these strains, we determined and examined the nucleotide sequences of the genomes of all available VRSA. Genome comparison identified candidate features that position strains of this lineage well for acquiring resistance to antibiotics in mixed infection.**

Received 13 April 2012 **Accepted** 24 April 2012 **Published** 22 May 2012

Citation Kos VN, et al. 2012. Comparative genomics of vancomycin-resistant *Staphylococcus aureus* strains and their positions within the clade most commonly associated with methicillin-resistant *S. aureus* hospital-acquired infection in the United States. mBio 3(3):e00112-12. doi:10.1128/mBio.00112-12.

Editor Paul Keim, Northern Arizona University

Address correspondence to Michael S. Gilmore, michael_gilmore@meei.harvard.edu.

Twelve cases of vancomycin-resistant *Staphylococcus aureus* (VRSA) infection have been reported in the United States since 2002 (1). Each is believed to represent a *de novo* acquisition of Tn*1546* from enterococci in a clonal cluster 5 (CC5) methicillin-resistant *S. aureus* (MRSA) (2). CC5 includes strains of pulsed-field gel electrophoresis (PFGE) types USA100 and USA800 and also contains the UK-EMRSA-3 strain, the New York-Japan clone, the Pediatric clone, the Rhine-Hesse epidemic strain, and the Canadian MRSA-2 strain (3).

CC5 strains are leading causes of hospital-associated *S. aureus* infection in the United States (4). They predominate in burn units, among blood isolates, and in intensive care nurseries (5–8) and rank among the leading causes of *S. aureus* infection globally (9, 10). CC5 strains were identified among early methicillin-resistant isolates in the 1960s (11) and were shown to have acquired staphylococcal cassette chromosome *mec* (SCC*mec*) at least 23 separate times (10). MRSA strains with reduced susceptibility to glycopeptide antibiotics (vancomycin- or glycopeptide-intermediate *S. aureus* [VISA or GISA, respectively]) (12) arise by spontaneous point mutations in cell wall synthesis genes (13) and are almost always CC5 (14).

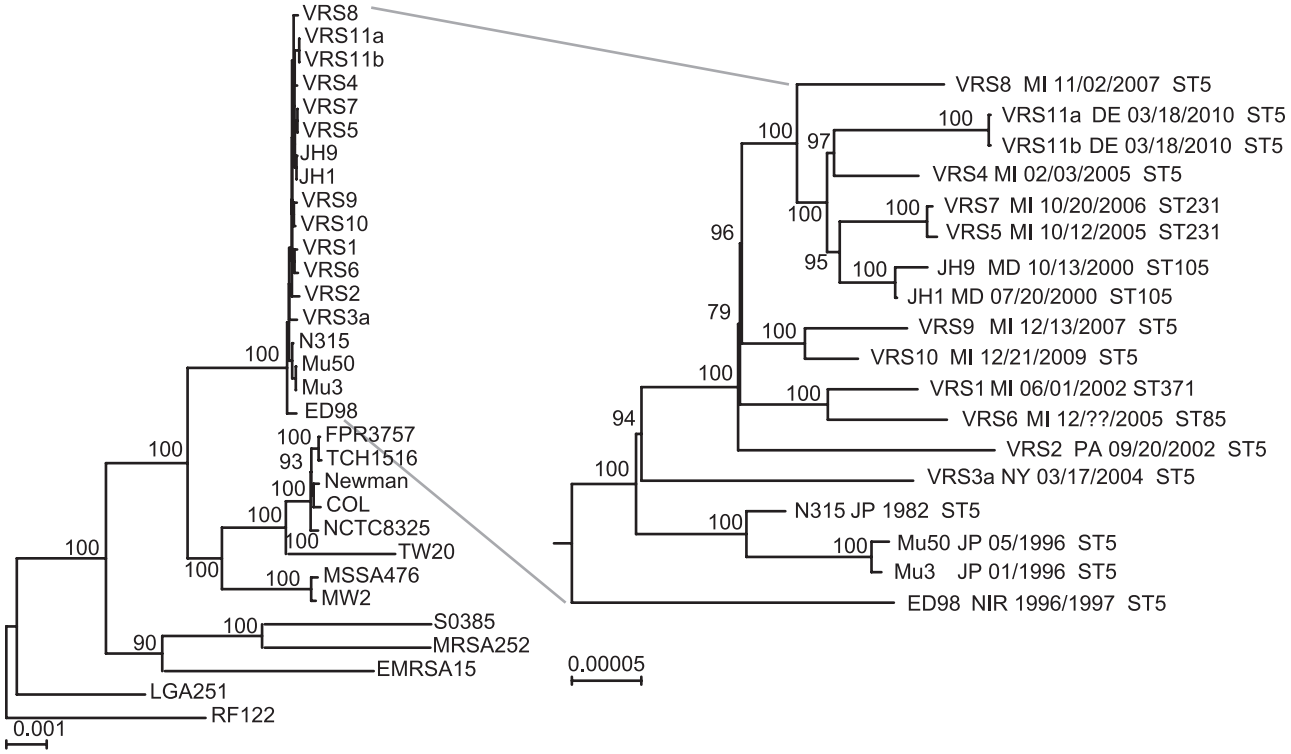

FIG 1 Phylogeny based on single-copy core orthologs (n = 1,822). Phylogeny showing the relationship of VRSA genomes to other completely sequenced *S. aureus* genomes. The date (month/day/year) and geographic location of the VRSA strain and multilocus sequence type (MLST) are indicated in the expanded section. MI, Michigan; PA, Pennsylvania; NY, New York; DE, Delaware; MD, Maryland; JP, Japan; NIR, Northern Ireland. Bootstrap values are indicated at each node. Scale bars correspond to number of changes per site.

Each of the 12 U.S. VRSA strains are believed to have resulted from acquisition of Tn*1546* from enterococci during the course of infection (15). Tn*1546* confers the VanA phenotype. Curiously, most of the Tn*1546* were from an *Enterococcus faecalis* donor (15–17), as opposed to the more common vancomycin-resistant *Enterococcus faecium* (18). Nine VRSA strains arose in southeast Michigan, where Tn*1546* transfer was mediated by a broad-host-range conjugative Inc18 plasmid (15–17, 19).

The repeated acquisition of vancomycin resistance by CC5, along with its involvement in the early acquisition of methicillin resistance and resistance to other antibiotics (11), suggests that it is genetically or biologically predisposed to horizontal acquisition of resistance and possibly other genes. Such transfer requires that donors and recipients coexist intimately in a mixed community, and that they achieve a population size that allows them to overcome inefficiencies and obstacles to transfer, genetic element establishment, and resistance expression. Thus, it was of interest to examine VRSA genomes for barriers to entry of foreign DNA as well as for traits that could foster their existence in mixed infection with potential resistance donors. We therefore generated high-quality draft genome sequences of the available 12 CC5 VRSA isolates from the first 11 VRSA cases in the United States and examined them for traits that may have predisposed this lineage to vancomycin resistance acquisition.

RESULTS

North American CC5 phylogeny. We determined a core gene sequence-based phylogeny for VRSA, based on 1,822 single-copy orthologs present in all genomes (Fig. 1). Strains do not cluster based on site or time of isolation (Fig. 1), supporting their independent development into VRSA. Strain VRS3a, isolated in New York in 2004 (the only PFGE type USA800 strain) is the most divergent. MRSA strain JH1 and its GISA derivative JH9, which arose during antimicrobial therapy (13, 20), are nested deeply within the VRSA. This phylogeny shows that all VRSA strains stem from a monophyletic source, supporting the hypothesis that they harbor a trait or traits that predispose them for vancomycin resistance acquisition or expression.

Tn*1546* haplotype network. To determine whether Tn*1546* evolved along a path different from the rest of the chromosome (as would be expected if Tn*1546* were repeatedly, independently acquired as opposed to having been acquired once and then passed along vertically with the rest of the chromosome), Tn*1546* sequences were compared, and their relationships to each other and to possible donor elements were calculated (17, 19) (Fig. 2). Tn*1546* sequences segregate regionally, as opposed to temporally (e.g., transposons from strains isolated in New York, Pennsylvania, and Delaware [strains VRS2, VRS3a, and VRS11a or VRS11b {VRS11a/b}]) share features with each other that are not shared by strains from Michigan (e.g., strains VRS1 and VRS4), consistent with an independent acquisition model (Fig. 2). All Tn*1546* sequences from Michigan VRSA strains form a tight cluster with few single-nucleotide polymorphisms (SNPs). Tn*1546* in *E. faecalis* VRE5, coisolated with VRS5 (17), possesses a 491-bp duplication missing in VRS5, suggesting that either VRE5 is not the Tn*1546* donor or that the duplication occurred in the donor after transfer.

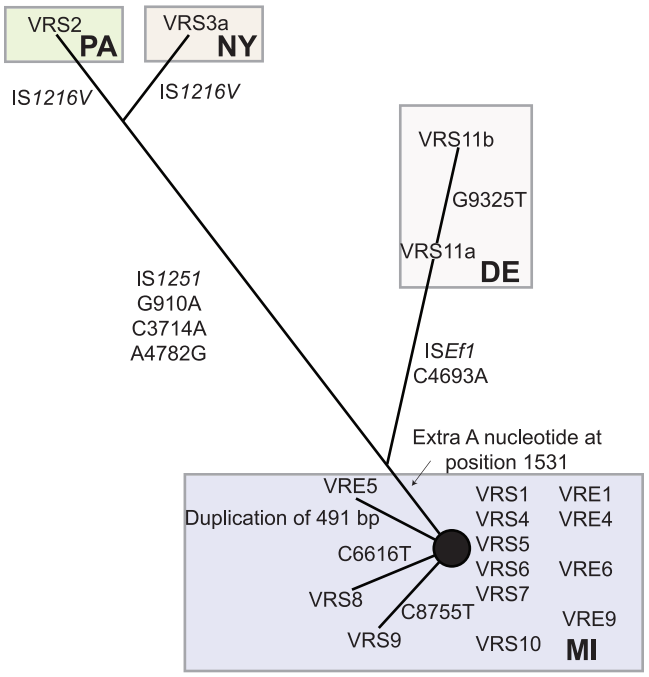

FIG 2 Haplotype network of Tn*1546* sequences. Numbering of the nucleotide changes refers to the position in sequence in comparison to the prototypical Tn*1546*(GenBank accession no. M97297).

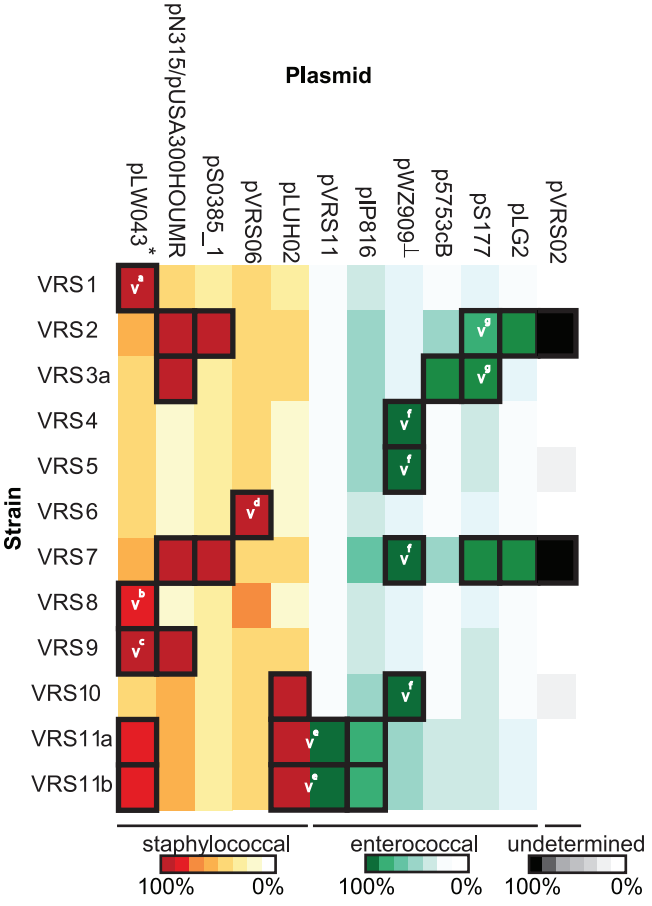

FIG 3 Plasmid sequences in VRSA. Heat map showing the extent of occurrence of known enterococcal (graded green shades) and staphylococcal plasmids (graded yellow to red) (based on detail in Table S5 in the supplemental material). Plasmids that are definitely present in the strains are indicated by squares outlined with a thick black border. The plasmid on which Tn*1546* resides is indicated by a white v, and superscript letters a through g denote unique insertion sites. pLW043* is the pSK41 homolog into which Tn*1546* inserted. pWZ909⊥ is the prototype Inc18 enterococcal plasmid involved in vancomycin resistance transfer. pN315 and pUSA300 were shown in one column, as these plasmids are nearly identical.

Tn*1546* of candidate enterococcal donor VRE4 is identical to that in VRS4, as well as to transposons in putative enterococcal donor strains VRE6 and VRS6. Tn*1546* from the plasmid in putative donor VRE9 differs from that in VRS9 by a single SNP (Fig. 2).

Tn*1546* of strains VRS2 and VRS3a were reported to possess insertions of IS*1251* between *vanS* and *vanH*, as well as copies of IS*1216* within *orf1* that are inverted in the two strains (21, 22). The Tn*1546* portions of these composite elements have 4 SNPs in common, distinguishing them from the VRS1 prototype and supporting their origination from a closely related source. Tn*1546* of strain VRS11a/b shares a SNP with strains VRS2 and VRS3a, placing it between the New York and Pennsylvania elements, and the Michigan transposons. VRS11a/b possesses an IS*E.faecalis1* (IS*Ef1*) insertion in the intergenic region between *vanX* and *vanY* (Fig. 2). VRS11a and VRS11b also possess a common SNP that is not found in other strains. Finally, even though isolated from a single patient, a third SNP distinguishes VRS11b from VRS11a, probably the result of microevolution within the patient.

Comparison of the Tn*1546* haplotype network to the single-copy core gene phylogeny showed that that there was little congruence and that the Tn*1546* haplotype network was a poor predictor of core gene phylogeny ($P < 0.0001$ by the Shimodaira-Hasegawa test [23]). Phylogenies based upon typing schemes for the hypervariable repeat regions of clumping factor B (*clfB*), coagulase (*coa*), and protein A (*spa*) and based on SCC*mec* and *agr* loci also failed to show a consistent line of VRSA strain descent (see Table S1A and Fig. S1A, S1B, and S1C in the supplemental material). The results of each analysis strongly support the model that each VRSA isolate arose as the result of an independent transposon acquisition.

Diversity in plasmid content and Tn*1546* location. Tn*1546* resides on plasmids in all VRSA strains, not on the chromosome (even though the chromosome is much larger and would seem to be a more probable insertion target). Each VRSA strain harbors plasmids of enterococcal and staphylococcal origin, and in some cases, cointegrates of the two (Fig. 3), highlighting their history of cooccurrence in mixed infections with enterococci. Plasmids involved in Tn*1546* acquisition vary. Strains VRS1 (24), VRS8, VRS9, and VRS11a/b possess large portions of the staphylococcal plasmid pSK41. pSK41 has been shown to enhance gene transfer from *E. faecalis* strains harboring pheromone-responsive plasmid pAD1 because of pheromone cross talk (25) and the pheromone encoded by a pSK41 gene is detected in the supernatant of strain VRS1 (15). The transposition of Tn*1546* onto pSK41, with concomitant loss of the enterococcal donor plasmid, was previously reported to have occurred in strain VRS1 (24). In strains VRS8 and VRS9, Tn*1546* also transposed to pSK41, but at unique sites. Strains VRS2, VRS3a, VRS4, VRS5, VRS6, VRS7, and VRS10 possess small fragments with sequence identity to pSK41, but these identities do not include pSK41 *traH*, the gene encoding the enterococcal plasmid pAD1 pheromone precursor.

Strains VRS2 and VRS3a acquired Tn*1546* on large (>100-kb) enterococcal plasmids, which were retained (21, 22). Identical sequences surround Tn*1546* in both strains, indicating a closely related source. These megaplasmids likely represent cointegrates of enterococcal plasmids pS177, p5753cB, and possibly pLG2 (Fig. 3). Mosaic plasmid structures are common in enterococci (26). Interestingly, the plasmid content of strain VRS7 is similar to that of strain VRS2, but with Tn*1546* remaining on the broad-host-range Inc18 donor plasmid (Fig. 3). In addition to VRS7, VRS4 and VRS5 also retain Tn*1546* on the Inc18 enterococcal donor plasmid (17). The sequence of the Inc18 plasmid in VRS10 was completed and found to possess an additional 2.97-kb transposon insertion not present in the others. In VRS4, VRS5, VRS7, and now VRS10, the location of Tn*1546* in the Inc18 donor plasmid is identical, highlighting the role of this common element in the Michigan outbreak (16). In VRS6, Tn*1546* was known to have transposed onto an *S. aureus* resident plasmid, with loss of the donor plasmid (16). In this strain, insertion occurred in a novel plasmid of apparent staphylococcal origin.

Strains VRS11a and VRS11b originated from a single patient in Delaware, and each strain possesses a cointegrate of enterococcal and staphylococcal plasmids (see Fig. S2 in the supplemental material) as confirmed by PCR. The cointegrate is a fusion of *S. aureus* plasmid pLUH02 (including genes for beta-lactamase, enterotoxin, cadmium resistance, and replication), and an enterococcal pCF10-like plasmid carrying pheromone-responsive genes. Tn*1546* resides on this fusion plasmid in a location that is identical in both strains (Fig. S2).

No functional lesions identifiable in restriction-modification systems of most VRSA strains. The Sau1 restriction system (27) and a type III-like restriction system (28) represent known barriers to foreign DNA entry. The Sau1 restriction system consists of an endonuclease, HsdR, encoded by a gene distant from two pairs of specificity and modification subunits, which occur on genomic islands *ν*Sa*α* and *ν*Sa*β* (27). In VRSA, most Sau1 restriction systems are intact (Fig. 4A). Strain VRS3a possesses a mutation in the *hsdR* Sau1 endonuclease gene that results in a large truncation of the HsdR subunit (see Fig. S2B in the supplemental material). Several polymorphisms in the *ν*Sa*α* *sau1CC5 hsdM1* copy of the HsdM modification subunit were observed (Fig. 4A), most representing minor amino acid changes. However, in the *ν*Sa*α* copy of HsdM in strain VRS9, a frameshift occurs, caused by an adenine duplication at nucleotide position 1213, which truncates the primary translation product (Fig. 4A). In the *ν*Sa*α* *sau1CC5 hsdS1* encoded specificity subunit of VRS11a/b, a nonsense mutation occurs in the second codon, likely causing a large truncation (Fig. 4A). In contrast to polymorphic copies in *ν*Sa*α*, only one polymorphism was noted in the *ν*Sa*β* copies of HdsS and HsdM subunits (Fig. 4A). That polymorphism corresponds to an addition of 3 amino acids to the carboxy terminus of HsdS in strain VRS2, which seems unlikely to alter function. Lack of mutations in the *ν*Sa*β* genes for HsdM and HsdS, as well as in most copies of counterparts in the *ν*Sa*β* island, indicate that Sau1 is likely functional in all VRSA strains, except for VRS3a.

A nonsense mutation in the type III-like restriction system has been reported to occur in Japanese CC5 strains (28). The impact of this specific mutation on function is unclear, although the type III-like restriction system clearly poses a barrier to DNA uptake (28). Only strain VRS10 possesses a polymorphism in this locus (MQS_01626; a point change that creates a nonsense mutation in the 12th codon) that is likely to be of functional consequence (Fig. 4A). Strain VRS9 possesses a polymorphism that leads to a conservative N89D amino acid substitution. Otherwise, the type III-like restriction systems of all other VRSA strains appear to be intact.

Frameshift in *dprA* in all VRSA strains except VRS3a. One coding difference related to DNA metabolism that stood out in comparison of CC5 to non-CC5 genomes was an adenine duplication at positions 333 and 334 in *dprA*. This duplication introduces a truncating frameshift (Fig. 4B), potentially eliminating or altering DprA function. This change was found in all North American CC5 strains except strain VRS3a. DprA influences DNA transformation efficiency in *Bacillus subtilis* (29). In addition to identifying potential loss-of-function mutations, such as the *dprA* frameshift, we also searched for other potentially function-altering nonsynonymous-codon-changing SNPs unique to North American CC5 strains. We identified 45 SNPs common to all except the most divergent strain, VRS3a (see Table S2 in the supplemental material).

VRSA strains possess polymorphisms in the *agr* locus. Similar to recent observations for phage 80/81 isolates of *S. aureus* that were prevalent in hospitals in the early 1990s (30), several VRSA strains (VRS1, VRS2, and VRS8 [see Fig. S1C in the supplemental material]) possess lesions in a key global regulator of virulence, *agr* (31). The *agr* locus is known to attenuate during infection (32), and this parallels reduced vancomycin susceptibility (33, 34).

CC5 gene content differences include lack of a bacteriocin operon and presence of genes for diverse superantigens. We next examined North American CC5 strains, including VRSA, for differences in gene content that could promote coexistence in polymicrobic infection (see Table S3A and S3B in the supplemental material). Most of the variable gene content unique to the North American CC5 strains occurs within the *ν*Sa*β* island (35, 36) (Table S3A; Fig. 5). Of potentially high importance, an operon (*bsa*) encoding a lantibiotic bacteriocin (active against other Gram-positive bacterial species [37, 38]), is absent in all members of the CC5 lineage. Instead, this island contains a unique cluster of enterotoxin genes (Fig. 5). Interestingly, this pattern—the absence of the bacteriocin operon and the presence of nearly identical complement of superantigens—also occurs in the *ν*Sa*β* pathogenicity islands in the EMRSA-15 and MRSA252/EMRSA-16 lineages that are prevalent in hospitals in the United Kingdom (39), even though the genetic backgrounds are highly dissimilar (Fig. 5). This suggests active selection for this version of *ν*Sa*β* in the hospital environment. The main difference in *ν*Sa*β* islands in strains from the United States and United Kingdom is a polymorphism that breaks the *seu* superantigen gene into *ψent1* and *ψent2* pseudogenes in the CC5 strains. In contrast to a previous hypothesis (40), our results indicate that the complete *seu* gene is in the ancestral state, and a deletion at the base of the CC5 clade created the pseudogenes. The North American CC5 *ν*SA*β* island also includes leukocidin genes, which encode a toxin that prevents phagocytic clearance (41, 42). CC5 strains lack the phage carrying the Panton-Valentine leukocidin (PVL) toxin (43) gene. Synteny analysis independently confirmed each of the above differences and identified other changes in gene position of unknown consequence (see Fig. S3 in the supplemental material).

Lipoproteins unique to CC5. *S. aureus* strains are known to harbor clusters of lipoprotein genes typically at four locations in the chromosome—within a *ν*Sa*α* element and at three other sites

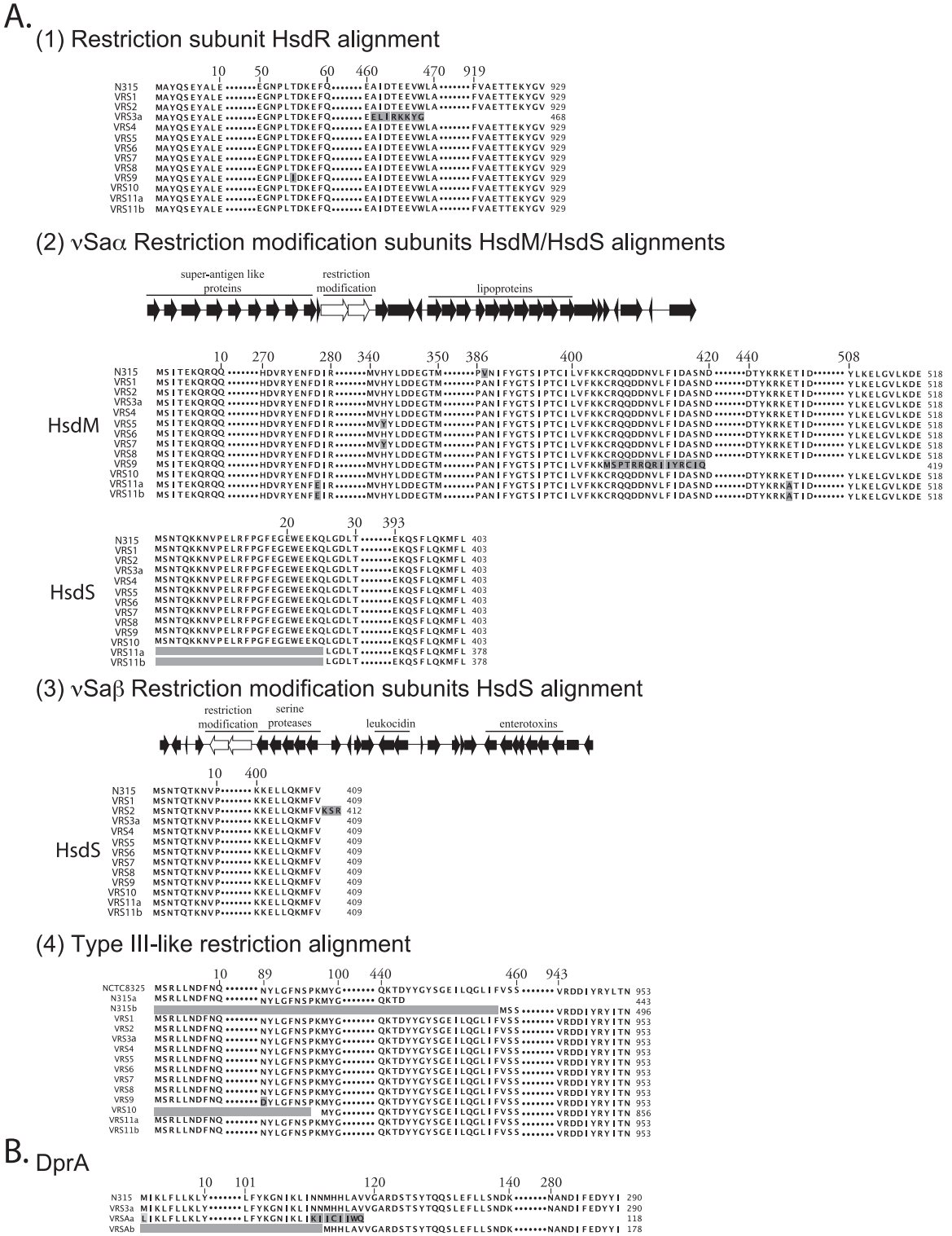

FIG 4 Restriction systems and *dprA* in VRSA. (A) (1) Alignment of inferred primary translation products of genes encoding the Sau1 restriction nuclease subunit (HsdR) compared to the CC5 strain originating in Japan, *S. aureus* N315. Polymorphisms are highlighted. Periods in blocks of sequence denote stretches that are identical and not shown. Numbers above the sequences represent the amino acid position in the prototype, from which distances represented by dots can be discerned. (2) Alignment of inferred primary translation products of genes encoding the modification (HsdM) and specificity (HsdS) subunits in the *ν*Sa*α* island. The locations of restriction-modification subunits relative to other key genes are indicated. (3) Alignment of inferred primary translation products of genes encoding the specificity (HsdS) subunits in the *ν*Sa*β* island. No polymorphisms occur in the *ν*Sa*β* HsdM-encoding gene (not shown). (4) Alignment of inferred primary translation products of genes encoding the type III-like restriction system compared to those of N315 and the CC8 *S. aureus* NCTC8325 prototype sequence. (B) Alignment of inferred primary translation products of *dprA* genes of VRSA. VRS3a is the only isolate that encodes a *dprA* product identical to that encoded by the Japanese CC5 isolate N315. All other VRSA strains possess a truncating frameshift mutation predicted to generate a premature termination, and possibly overlapping reinitiation product, as generically designated in the figure as VRSAa and VRSAb.

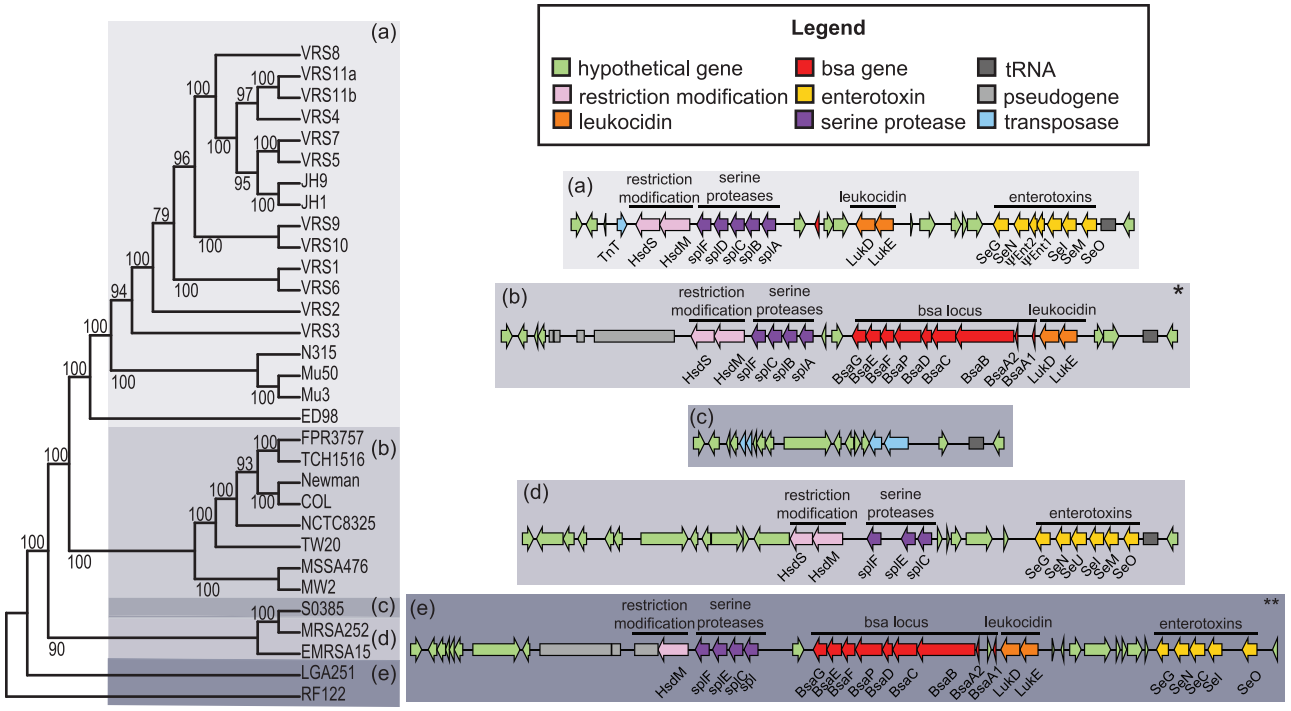

FIG 5 Variation in the *ν*SA*β* island of VRSA strains compared to other lineages. The gray shading for the schematic representations of the *ν*SA*β* islands shown to the right of the figure corresponds to the position within the phylogeny showing the relationship of VRSA genomes to other completely sequenced *S. aureus* genomes shown to the left of the figure. The asterisk in the top right-hand corner of the schematic representation in panel b indicates that *S. aureus* COL contains an IS*1811* transposase upstream of *bsaA2*. The pair of asterisks in the top right-hand corner of the schematic representation in panel d indicate that the far right endpoint for RF122 is not shown because of phage insertion.

(44). CC5 genomes possess a significantly larger set of lipoprotein genes than non-CC5 genomes ($P < 0.05$ by the Mann-Whitney U test; see Fig. S4 in the supplemental material), suggesting that selection favors their occurrence in these hospital isolates. We observed strong congruence between lipoprotein- and whole-genome-based phylogenies, showing that lipoprotein variation largely parallels chromosome divergence patterns. Interestingly, BLAST identified *Listeria grayii* as the only nonstaphylococcal species possessing highly related homologs of *S. aureus* lipoproteins (<1e−24). Gram-positive bacteria process lipoprotein signal peptides into septa- or octapeptide pheromones (45), which have been hypothesized to contribute acquisition of vancomycin resistance by staphylococci (46). Potential pheromone sequences encoded by the North American CC5 *S. aureus* genomes were identified (Fig. S4).

DISCUSSION

MRSA emerged in the early 1960s and remained largely restricted to the hospital environment (43). Not until resistance occurred in other lineages 30 years later did it rapidly spread in the community (43). CC5 is often the first lineage in which new antibiotic resistance genes appear (10, 11, 47). Our work provides quantitative support for an earlier proposition, based largely on PFGE (15–17), that each occurrence of Tn*1546*-conferred vancomycin resistance in *S. aureus* represents an independent acquisition, rather than patient-to-patient spread. Modeling evolutionary distances on a time scale (Fig. 6) shows that the last common ancestor of all VRSA strains occurred at about the time of methicillin introduction, about 1960, 40 years earlier than vancomycin resistance was found to have entered the species. The early 1960s also is the approximate time when the Japanese CC5 isolates diverged. The modeling analysis employed was based on a calculated mutation rate of 3.46×10^{-6} per site per year and employed conservative, relaxed-clock assumptions. This rate is in excellent agreement with the $3.3 \times 10^{-6} \pm 0.7 \times 10^{-6}$ mutation rate calculated independently by others in a rigorous application of second-generation sequencing in a study of the global spread of the MRSA lineage ST239 that is prominent in Asia (48). Only in the case of strains VRS11a and VRS11b, which were isolated from the same patient, does the uncertainty with respect to the time of strain divergence extend beyond the date of isolation of that VRSA strain. The 1960 time point represents the time when the most divergent VRSA strain, VRS3b (PFGE type USA800), branched from the rest, most being USA100 PFGE type (Fig. 5). Within the USA100 group, extensive diversification occurred next in about 1978.

In contrast to expectations, VRSA restriction barriers appear to be largely intact. Strain VRS3a possesses a defect in the Sau1 endonuclease that is likely to be of functional consequence. However, most other polymorphisms in the Sau1 system were limited to the *ν*Sa*α*-encoded copy of a modification gene, with the *ν*Sa*β*-encoded copy fully intact. Another strain, VRS10, possesses change in the type III-like restriction, shortening the predicted primary translation product from 953 amino acids to 856 amino acids by removal of the amino terminus. The functional consequence of this truncation is currently unknown. It may be important that all VRSA strains, except for the phylogenetic outlier strain VRS3a, possess a nonsense mutation early in the *dprA* gene

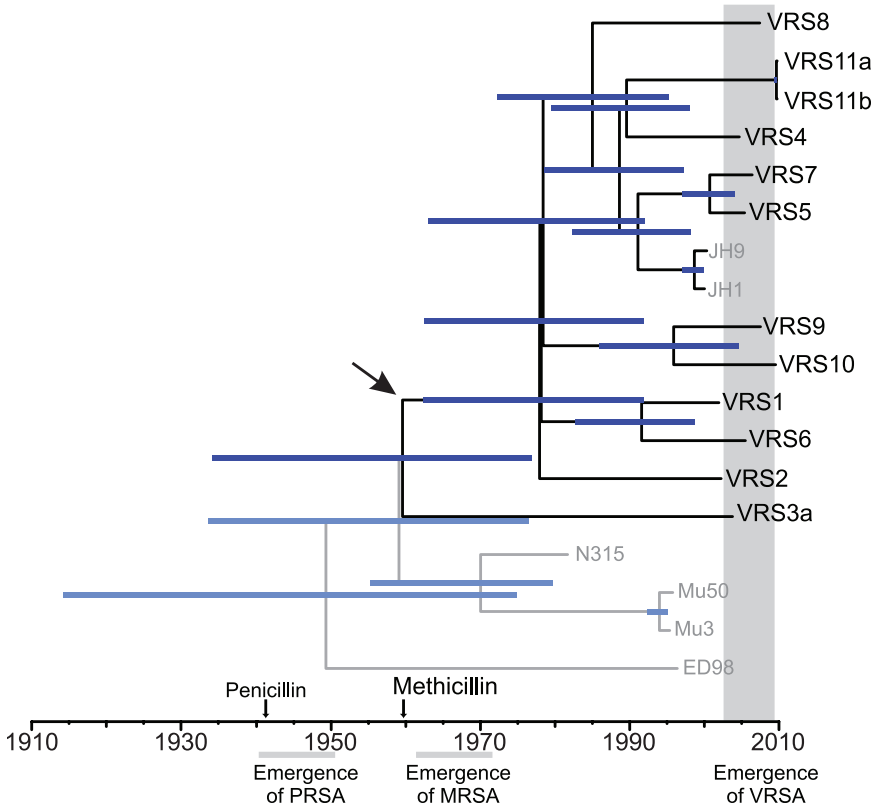

FIG 6 Consensus tree for CC5 under relaxed clock conditions. North American CC5 strains are shown in black, while other strains are shown in gray. Blue bars indicate range of 95% highest posterior density interval (95% confidence interval, 1.76×10^{-6} to 5.22×10^{-6}). The time of penicillin and methicillin introduction, as well as the emergence of resistances are shown for reference (PRSA, penicillin-resistant *S. aureus*). The arrow indicates the estimated time of divergence of the most distantly related VRSA genome, that of VRS3a (USA800 branch) from the remaining VRSA (mainly USA100) branch of the CC5 clade, showing that strain divergence occurred far earlier than vancomycin resistance acquisition, supporting the model of independent Tn*1546* acquisition.

that is predicted to truncate a majority of the polypeptide. DprA (also known as Smf [49]), is highly conserved and contributes to efficient DNA transformation in naturally competent bacteria (29, 49–54). Transformation efficiency of plasmids in a *B. subtilis dprA* mutant is decreased 60-fold (29). Experiments with *Escherichia coli dprA* mutants do not show an obvious role in transformation or conjugation (55). Its function in *S. aureus* remains to be explored.

The most variable feature of the VRSA genome is plasmid content. In all cases, Tn*1546* resides on a plasmid, even though it clearly transposed upon entry into some strains, and because of size, the chromosome would seem to be the most probable target for transposon insertion. The basis for the insertion site preference for plasmids over the *S. aureus* chromosome, and also for an apparent incompatibility between the enterococcal Inc18 plasmid that played a major role in the Michigan outbreak and an endogenous *S. aureus* pSK41 plasmid present in several recipients, is unknown. VRSA genomes are replete with plasmids of enterococcal origin, highlighting their cooccurrence in polymicrobic infections and possibly in other ecologies. The multiplicity of plasmid structures conveying Tn*1546*, including *S. aureus*/enterococcal cointegrate plasmids, increases the odds of future transfers, possibly into staphylococcal lineages or species where a lower fitness cost is incurred.

The genomic island νSaβ is a distinguishing feature of CC5. Of potential ecological importance, the *bsa* operon usually encoded within this island is absent in VRSA and other CC5 strains. Frequent application of antibiotics in the hospital environment may select for strains with an enhanced ability to comingle with potential resistance donors of other species. Interestingly, this trait is also lacking in otherwise highly divergent strains that are prevalent in hospitals in the United Kingdom. The fact that United Kingdom and U.S. hospital strains with widely different chromosomal backgrounds share very similar νSaβ islands suggests that there may be active selection for this configuration in the hospital environment. Multidrug-resistant enterococci are much more likely to lack clustered regularly interspaced short palindromic repeat (CRISPR) defenses of the genome than commensal strains (56), indicating that the widespread use of antibiotics has selected for hospital-adapted bacteria that have enhanced abilities to exist in mixed communities and exchange resistance determinants. Loss of bacteriocin production as well as immunity may also explain why, for 30 years, CC5 MRSA strains were not able to establish methicillin resistance in the community at a high level. They may have been inhibited by the functional bacteriocin loci of *S. aureus* strains of other sequence types (such as CC8 and CC1) already present in the community niche. This may also be limiting the spread of vancomycin resistance from CC5 strains to other clades.

Staphylococcal enterotoxins, T cell mitogen superantigens (57) that dysregulate the host response by stimulating $CD8^+$ regulatory T cells (T_{regs}) at low concentrations (58) and other mechanisms, are particularly abundant in the CC5 νSaβ element, as well as that of the

United Kingdom clones that are prevalent in hospitals. Enterotoxins, together with lipoproteins (59) and leukocidin (41, 42) may facilitate overgrowth of mixed populations of bacteria at a site of infection, contributing to the creation of conditions favorable for resistance transfer. Notably, most VRSA strains occurred in mixed infections of plantar ulcers of diabetic patients (2), infections that are known to be highly polymicrobic (60, 61).

The hospital is a unique environment, where colonization and patient-to-patient propagation of a strain may depend less on bacterial virulence traits associated with transmission (62) than on transmission vectors in the form of hospital staff and environmental surfaces and antibiotics (63). Strains that are prevalent in hospitals are under continuous antibiotic selection pressure and are exposed to an ever rotating arsenal. CC5 isolates appear to be very well adapted for surviving and evolving in this environment by acquiring resistance to new antibiotics.

MATERIALS AND METHODS

Strains. All available VRSA strains were obtained through the Network of Antimicrobial Resistance in *Staphylococcus aureus* (NARSA) (see Table S1B in the supplemental material). Since strain VRS3a is believed to be identical to strain VRS3b, the latter was not sequenced. Strains VRS11b and VRS11a had not been characterized, and both were examined because of differences in oxacillin resistance. The VRSA strains were routinely grown on tryptic soy agar containing 10 μg/ml vancomycin.

Genome sequencing. For Illumina sequencing, total DNA was purified from 10-ml overnight cultures using the DNeasy DNA extraction kit (Qiagen). DNA was transferred to the Tufts University DNA Core Facility, Boston, MA, and a modified Illumina protocol (64) was used. Libraries were subjected to multiplexed paired-end sequencing according to the manufacturer's specifications. Sequencing reads were filtered to exclude reads with a quality score of <25 at any position. The average coverage of the 3-Mb genomes was >110-fold. The genomes were also independently sequenced at the University at Buffalo Next-Generation Sequencing and Expression Analysis Core (Buffalo, NY), by 454 FLX (Roche) to at least 10-fold coverage.

Assemblies. Illumina reads were assembled using Velvet version 1.0.18 (65). The 454 and Illumina reads were then combined and assembled using Newbler 2.3 (Roche). Gene annotations were generated using the Prodigal gene caller (66). Draft genomes for strains were submitted to GenBank under the following accession numbers: strain VRS1, AHBK00000000; VRS2, AHBL00000000; VRS3a, AHBM00000000; VRS4, AHBN00000000; VRS5, AHBO00000000; VRS6, AHBP00000000; VRS7, AHBQ00000000; VRS8, AHBR00000000; VRS9, AHBS00000000; VRS10, AHBT00000000; VRS11a, AHBU00000000; and VRS11b, AHBV00000000. The complete genomes of *S. aureus* strains used for comparison (see Table S1C in the supplemental material) were downloaded from GenBank (http://www.ncbi.nlm.nih.gov) and the Sanger website (http://www.sanger.ac.uk/pathogens).

PCR and targeted DNA sequencing. To verify results in some cases or to obtain missing sequence information, PCR was performed using *Taq* polymerase (New England Biolabs) with the amplification primers listed in Table S4 in the supplemental material. DNA sequencing of individual PCR products was performed by the Massachusetts General Hospital DNA Sequencing Core Facility. Specific nucleotide sequence alignments were routinely performed using mafft (67) and ClustalW (68).

Bioinformatic analysis. See Text S1 in the supplemental material.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00112-12/-/DCSupplemental.

Text S1, DOC file, 0.2 MB.
Figure S1, EPS file, 5.7 MB.
Figure S2, EPS file, 0.4 MB.
Figure S3, PDF file, 6.8 MB.
Figure S4, PDF file, 1.5 MB.
Table S1, DOC file, 0.1 MB.
Table S2, DOCX file, 0.1 MB.
Table S3, DOC file, 0.1 MB.
Table S4, DOC file, 0.1 MB.
Table S5, DOC file, 0.1 MB.

ACKNOWLEDGMENTS

This research was supported in part with federal funds from the NIAID/NIH/DHHS, including the Harvard-wide Program on Antibiotic Resistance (NIH grant AI083214), NIH grant EY017381 (M.S.G.), contract HHSN27220090018C, and the NARSA program supported under NIAID/NIH contract HHSN272200700055C.

We thank Joseph J. Ferretti, Javier Irazoqui, and Ashlee Earl for helpful comments during manuscript preparation. We also thank Don B. Clewell, Jean Patel, Wenming Zhu, and Susan E. Flannagan for helpful discussions during the course of these studies.

REFERENCES

1. **CDC.** 2002. *Staphylococcus aureus* resistant to vancomycin—United States. Morb. Mortal. Wkly. Rep. **51**:565–567.
2. **Sievert DM, et al.** 2008. Vancomycin-resistant *Staphylococcus aureus* in the United States, 2002–2006. Clin. Infect. Dis. **46**:668–674.
3. **Monecke S, et al.** 2011. A field guide to pandemic, epidemic and sporadic clones of methicillin-resistant *Staphylococcus aureus*. PLoS One **6**:e17936.
4. **Klevens RM, et al.** 2007. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. JAMA **298**:1763–1771.
5. **Chua T, et al.** 2008. Molecular epidemiology of methicillin-resistant *Staphylococcus aureus* bloodstream isolates in urban Detroit. J. Clin. Microbiol. **46**:2345–2352.
6. **Limbago B, et al.** 2009. Characterization of methicillin-resistant *Staphylococcus aureus* isolates collected in 2005 and 2006 from patients with invasive disease: a population-based analysis. J. Clin. Microbiol. **47**:1344–1351.
7. **Murray CK, et al.** 2009. Twenty-five year epidemiology of invasive methicillin-resistant *Staphylococcus aureus* (MRSA) isolates recovered at a burn center. Burns **35**:1112–1117.
8. **Seybold U, et al.** 2008. Emergence of and risk factors for methicillin-resistant *Staphylococcus aureus* of community origin in intensive care nurseries. Pediatrics **122**:1039–1046.
9. **Deurenberg RH, Stobberingh EE.** 2008. The evolution of *Staphylococcus aureus*. Infect. Genet. Evol. **8**:747–763.
10. **Nübel U, et al.** 2008. Frequent emergence and limited geographic dispersal of methicillin-resistant *Staphylococcus aureus*. Proc. Natl. Acad. Sci. U. S. A. **105**:14130–14135.
11. **Crisóstomo MI, et al.** 2001. The evolution of methicillin resistance in *Staphylococcus aureus*: similarity of genetic backgrounds in historically early methicillin-susceptible and -resistant isolates and contemporary epidemic clones. Proc. Natl. Acad. Sci. U. S. A. **98**:9865–9870.
12. **CLSI.** 2006. Investigation and control of vancomycin-intermediate and -resistant *Staphylococcus aureus* (VISA/VRSA). A guide for health departments and infection control personnel. CLSI, Wayne, PA.
13. **Mwangi MM, et al.** 2007. Tracking the *in vivo* evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. Proc. Natl. Acad. Sci. U. S. A. **104**:9451–9456.
14. **McDougal LK, et al.** 2003. Pulsed-field gel electrophoresis typing of oxacillin-resistant *Staphylococcus aureus* isolates from the United States: establishing a national database. J. Clin. Microbiol. **41**:5113–5120.
15. **Flannagan SE, et al.** 2003. Plasmid content of a vancomycin-resistant *Enterococcus faecalis* isolate from a patient also colonized by *Staphylococcus aureus* with a VanA phenotype. Antimicrob. Agents Chemother. **47**:3954–3959.
16. **Zhu W, et al.** 2008. Vancomycin-resistant *Staphylococcus aureus* isolates associated with Inc18-like *vanA* plasmids in Michigan. Antimicrob. Agents Chemother. **52**:452–457.
17. **Zhu W, et al.** 2010. Dissemination of an *Enterococcus* Inc18-like *vanA* plasmid associated with vancomycin-resistant *Staphylococcus aureus*. Antimicrob. Agents Chemother. **54**:4314–4320.
18. **Huycke MM, Sahm DF, Gilmore MS.** 1998. Multiple-drug resistant enterococci: the nature of the problem and an agenda for the future. Emerg. Infect. Dis. **4**:239–249.

19. **Palmer KL, et al.** 2010. High-quality draft genome sequences of 28 *Enterococcus* sp. isolates. J. Bacteriol. **192**:2469–2470.
20. **Sieradzki K, Leski T, Dick J, Borio L, Tomasz A.** 2003. Evolution of a vancomycin-intermediate *Staphylococcus aureus* strain *in vivo*: multiple changes in the antibiotic resistance phenotypes of a single lineage of methicillin-resistant *S. aureus* under the impact of antibiotics administered for chemotherapy. J. Clin. Microbiol. **41**:1687–1693.
21. **Tenover FC, et al.** 2004. Vancomycin-resistant *Staphylococcus aureus* isolate from a patient in Pennsylvania. Antimicrob. Agents Chemother. **48**:275–280.
22. **Weigel LM, et al.** 2007. High-level vancomycin-resistant *Staphylococcus aureus* isolates associated with a polymicrobial biofilm. Antimicrob. Agents Chemother. **51**:231–238.
23. **Shimodaira H, Hasegawa M.** 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. **16**:1114–1116.
24. **Weigel LM, et al.** 2003. Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. Science **302**:1569–1571.
25. **Firth N, Fink PD, Johnson L, Skurray RA.** 1994. A lipoprotein signal peptide encoded by the staphylococcal conjugative plasmid pSK41 exhibits an activity resembling that of *Enterococcus faecalis* pheromone cAD1. J. Bacteriol. **176**:5871–5873.
26. **Sletvold H, et al.** 2010. Tn1546 is part of a larger plasmid-encoded genetic unit horizontally disseminated among clonal *Enterococcus faecium* lineages. J. Antimicrob. Chemother. **65**:1894–1906.
27. **Waldron DE, Lindsay JA.** 2006. Sau1: a novel lineage-specific type I restriction-modification system that blocks horizontal gene transfer into *Staphylococcus aureus* and between *S. aureus* isolates of different lineages. J. Bacteriol. **188**:5578–5585.
28. **Corvaglia AR, et al.** 2010. A type III-like restriction endonuclease functions as a major barrier to horizontal gene transfer in clinical *Staphylococcus aureus* strains. Proc. Natl. Acad. Sci. U. S. A. **107**:11954–11958.
29. **Tadesse S, Graumann PL.** 2007. DprA/Smf protein localizes at the DNA uptake machinery in competent *Bacillus subtilis* cells. BMC Microbiol. **7**:105.
30. **Deleo FR, et al.** 2011. Molecular differentiation of historic phage-type 80/81 and contemporary epidemic *Staphylococcus aureus*. Proc. Natl. Acad. Sci. U. S. A. **108**:18091–18096.
31. **Novick RP, Geisinger E.** 2008. Quorum sensing in staphylococci. Annu. Rev. Genet. **42**:541–564.
32. **Traber KE, et al.** 2008. *agr* function in clinical *Staphylococcus aureus* isolates. Microbiology **154**:2265–2274.
33. **Sakoulas G, et al.** 2003. *Staphylococcus aureus* accessory gene regulator (*agr*) group II: is there a relationship to the development of intermediate-level glycopeptide resistance? J. Infect. Dis. **187**:929–938.
34. **Sakoulas G, et al.** 2002. Accessory gene regulator (*agr*) locus in geographically diverse *Staphylococcus aureus* isolates with reduced susceptibility to vancomycin. Antimicrob. Agents Chemother. **46**:1492–1502.
35. **Baba T, Bae T, Schneewind O, Takeuchi F, Hiramatsu K.** 2008. Genome sequence of *Staphylococcus aureus* strain Newman and comparative analysis of staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands. J. Bacteriol. **190**:300–310.
36. **Fitzgerald JR, Sturdevant DE, Mackie SM, Gill SR, Musser JM.** 2001. Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. Proc. Natl. Acad. Sci. U. S. A. **98**:8821–8826.
37. **Daly KM, et al.** 2010. Production of the Bsa lantibiotic by community-acquired *Staphylococcus aureus* strains. J. Bacteriol. **192**:1131–1142.
38. **Scott JC, Sahl HG, Carne A, Tagg JR.** 1992. Lantibiotic-mediated antilactobacillus activity of a vaginal *Staphylococcus aureus* isolate. FEMS Microbiol. Lett. **72**:97–102.
39. **Holden MT, et al.** 2004. Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. Proc. Natl. Acad. Sci. U. S. A. **101**:9786–9791.
40. **Letertre C, Perelle S, Dilasser F, Fach P.** 2003. Identification of a new putative enterotoxin SEU encoded by the *egc* cluster of *Staphylococcus aureus*. J. Appl. Microbiol. **95**:38–43.
41. **Gravet A, et al.** 1998. Characterization of a novel structural member, LukE-LukD, of the bi-component staphylococcal leucotoxins family. FEBS Lett. **436**:202–208.
42. **Prevost G, et al.** 1998. Pore-forming leukotoxins from *Staphylococcus aureus*: variability of the target cells and 2 pharmacological processes. Pathol. Biol. (Paris) **46**:435–441. (In French).
43. **David MZ, Daum RS.** 2010. Community-associated methicillin-resistant *Staphylococcus aureus*: epidemiology and clinical consequences of an emerging epidemic. Clin. Microbiol. Rev. **23**:616–687.
44. **Tsuru T, Kobayashi I.** 2008. Multiple genome comparison within a bacterial species reveals a unit of evolution spanning two adjacent genes in a tandem paralog cluster. Mol. Biol. Evol. **25**:2457–2473.
45. **An FY, Sulavik MC, Clewell DB.** 1999. Identification and characterization of a determinant (eep) on the *Enterococcus faecalis* chromosome that is involved in production of the peptide sex pheromone cAD1. J. Bacteriol. **181**:5915–5921.
46. **Clewell DB, Francia MV, Flannagan SE, An FY.** 2002. Enterococcal plasmid transfer: sex pheromones, transfer origins, relaxases, and the *Staphylococcus aureus* issue. Plasmid **48**:193–201.
47. **Hiramatsu K, et al.** 1997. Dissemination in Japanese hospitals of strains of *Staphylococcus aureus* heterogeneously resistant to vancomycin. Lancet **350**:1670–1673.
48. **Harris SR, et al.** 2010. Evolution of MRSA during hospital transmission and intercontinental spread. Science **327**:469–474.
49. **Ogura M, et al.** 2002. Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. J. Bacteriol. **184**:2344–2351.
50. **Ando T, Israel DA, Kusugami K, Blaser MJ.** 1999. HP0333, a member of the *dprA* family, is involved in natural transformation in *Helicobacter pylori*. J. Bacteriol. **181**:5572–5580.
51. **Bergé M, Mortier-Barrière I, Martin B, Claverys JP.** 2003. Transformation of *Streptococcus pneumoniae* relies on DprA- and RecA-dependent protection of incoming DNA single strands. Mol. Microbiol. **50**:527–536.
52. **Friedrich A, Prust C, Hartsch T, Henne A, Averhoff B.** 2002. Molecular analyses of the natural transformation machinery and identification of pilus structures in the extremely thermophilic bacterium *Thermus thermophilus* strain HB27. Appl. Environ. Microbiol. **68**:745–755.
53. **Karudapuram S, Zhao X, Barcak GJ.** 1995. DNA sequence and characterization of *Haemophilus influenzae dprA+*, a gene required for chromosomal but not plasmid DNA transformation. J. Bacteriol. **177**:3235–3240.
54. **Takata T, Ando T, Israel DA, Wassenaar TM, Blaser MJ.** 2005. Role of dprA in transformation of *Campylobacter jejuni*. FEMS Microbiol. Lett. **252**:161–168.
55. **Smeets LC, et al.** 2006. Functional characterization of the competence protein DprA/Smf in *Escherichia coli*. FEMS Microbiol. Lett. **263**:223–228.
56. **Palmer KL, Gilmore MS.** 2011. Multidrug-resistant enterococci lack CRISPR-*cas*. mBio **1**(4):e00227–10.
57. **Proft T, Fraser JD.** 2003. Bacterial superantigens. Clin. Exp. Immunol. **133**:299–306.
58. **Taylor AL, Cross EL, Llewelyn MJ.** 2012. Induction of contact-dependent CD8(+) regulatory T cells through stimulation with staphylococcal and streptococcal superantigens. Immunology **135**:158–167.
59. **Schmaler M, Jann NJ, Götz F, Landmann R.** 2010. Staphylococcal lipoproteins and their role in bacterial survival in mice. Int. J. Med. Microbiol. **300**:155–160.
60. **Dowd SE, et al.** 2008. Polymicrobial nature of chronic diabetic foot ulcer biofilm infections determined using bacterial tag encoded FLX amplicon pyrosequencing (bTEFAP). PLoS One **3**:e3326.
61. **Grice EA, et al.** 2010. Longitudinal shift in diabetic wound microbiota correlates with prolonged skin defense response. Proc. Natl. Acad. Sci. U. S. A. **107**:14799–14804.
62. **Massey RC, Horsburgh MJ, Lina G, Höök M, Recker M.** 2006. The evolution and maintenance of virulence in *Staphylococcus aureus*: a role for host-to-host transmission? Nat. Rev. Microbiol. **4**:953–958.
63. **Collins J, et al.** 2010. Offsetting virulence and antibiotic resistance costs by MRSA. ISME J. **4**:577–584.
64. **Palmer KL, Daniel A, Hardy C, Silverman J, Gilmore MS.** 2011. Genetic basis for daptomycin resistance in enterococci. Antimicrob. Agents Chemother. **55**:3345–3356.
65. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.
66. **Hyatt D, et al.** 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform. **11**:119.
67. **Katoh K, Misawa K, Kuma K, Miyata T.** 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. **30**:3059–3066.
68. **Larkin MA, et al.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23**:2947–2948.